

AD-A096 687

RICE UNIV HOUSTON TX DEPT OF PSYCHOLOGY

F/G 5/10

UNCERTAINTY MEASUREMENT IN A COMPLEX TASK AS A FUNCTION OF RESP--ETC(U)

FEB 81 W C HOWELL, S P KERKAR

N00014-78-C-0555

UNCLASSIFIED

TR-81-1

NL

[ or ]

42A  
736A97




END  
DATE  
FILMED  
4-81  
DTIC

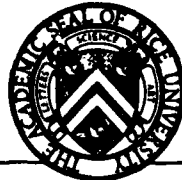
7c

2

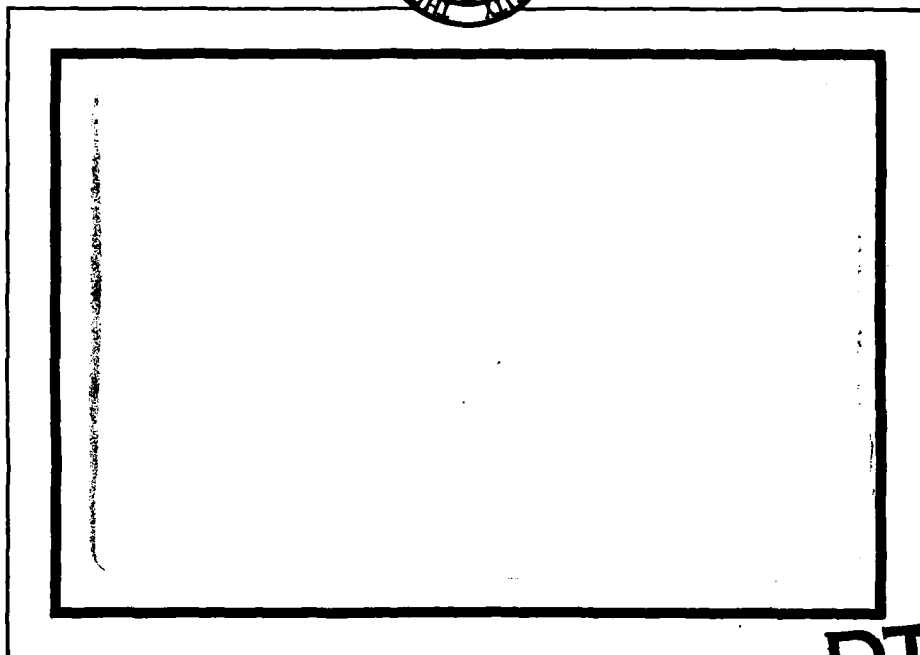
# RICE UNIVERSITY

Houston, Texas 77001

LEVEL II



AD A 096687



DTIC  
ELECTE  
MAR 23 1981  
S D E

Department of Psychology  
Research Report Series

FILE COPY

81 3 20 013

(2)

LEVEL II

(9) Technical rept.

16 Apr 81 - 15 Jan 81

(6)  
UNCERTAINTY MEASUREMENT IN A  
COMPLEX TASK AS A FUNCTION OF  
RESPONSE MODE AND EVENT TYPE  
CHARACTERISTICS.

(10)  
William C. Howell and Shanta P. Kerkar  
Rice University

Technical Report #81-1

(11) February 1981

15341

DTIC

This research was supported by the Engineering Psychology  
Programs, Office of Naval Research, ONR Contract N00014-78-C-0555  
Work Unit NR197-050.

(15)  
Approved for public release; distribution unlimited.

Reproduction in whole or part is permitted for any purpose of  
the United States government.

411654

13

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 81-1	2. GOVT ACCESSION NO. AD-A096687	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Uncertainty Measurement in a Complex Task is a Function of Response Mode and Event Type Characteristics		5. TYPE OF REPORT & PERIOD COVERED 4/16/80 - 1/15/81
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William C. Howell and Shanta P. Kerkar		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0555
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Rice University Houston, TX 77001		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR197-050
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) uncertainty judgment, choice behavior, probability estimation, frequency estimation, calibration, task effects, predecision processes		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) (see <u>abstract</u> )		

Abstract

This research was concerned with the hypothesized effect of response mode and event type factors on measured uncertainty for frequentistic events observed within a realistic task setting. Subjects served as dispatchers of emergency services for a hypothetical city on five daily shifts. Events were generated by a stationary stochastic process (emergency calls) or by the subject's responses to these calls (allocations and correct responses). At the conclusion of the 600 dispatching trials, subjects in the two experimental groups estimated either the frequency (FE) or probability (PE) of various kinds of events, and in a subsequent session, made predictive choices among selected event pairs. A control group (C) followed the same procedure except for omission of the estimation task. Results showed that estimation performance is influenced reliably by both variables: the FE group was superior to the PE group on all events, externally generated events produced generally better estimates than "internal" ones, and spatial events were judged more accurately than non-spatial ones. Experimental groups made better choices than the control group, and there was a tendency for choice performance to reflect the quality of estimation performance. Possible explanations for the influence of estimation on choice measures are discussed together with various practical, theoretical, and research implications.

Unannounced Justification		<input checked="checked" type="checkbox"/>
By _____		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A		

## INTRODUCTION

Since few behavioral contexts afford people the luxury of complete, dependable information, any attempt to understand cognitively directed behavior must deal with the intuitive processing of uncertainty. This point has become widely recognized and has stimulated descriptive research on a variety of issues related to the formation or use of probabilistic concepts (Einhorn & Hogarth, 1981; Hammond, McClelland, & Mumpower, 1980; Slovic, Fischhoff, & Lichtenstein, 1977; Wallsten, 1980). Attempts to model or "capture" the policy of decision makers (Dawes & Corrigan, 1974; Naylor & Domine, 1981), to "calibrate" weather forecasters (Lichtenstein, Fischhoff, & Phillips, 1977), to track or model the way people revise opinion in light of evidence (Slovic & Lichtenstein, 1971), to explain why feedback is often ineffective (Einhorn & Hogarth, 1981), and to describe how people store event frequency information (Marques & Howell, Note 2), are but a few illustrations of this activity.

Clearly the most influential work of the past decade was the demonstration by Tversky, Kahneman, and their coworkers that many of the axioms of probability theory are inappropriate for the description of human information processing; that people operate according to convenient "heuristics" rather than "rational" strategies in the classical sense and hence are subject to systematic "biases" (Kahneman & Tversky, 1973; Tversky & Kahneman, 1981). This line of thinking has shifted the emphasis of research from the normative to the descriptive and has raised the possibility that task demands may play a vital

role in determining how people deal with uncertainty in any particular situation (Slovic, Fischhoff, & Lichtenstein, 1977). In support of this view, Tversky & Kahneman (1981) have shown that formally identical problems can produce overwhelmingly different choice preferences depending upon the way in which the uncertainty is "framed" or expressed. Cohen (1979) has also emphasized the importance of the way decision problems are formulated, but does not regard the resulting shifts in behavior as an indication of human deficiencies or biases in inference.

In view of the diversity of events, task settings, response requirements, and conditions under which uncertainty has been measured, the first step toward a better understanding of task effects would seem to be identification and classification of key task parameters. Several attempts at developing such taxonomies have, in fact, appeared in recent years (Hammond, McClelland, & Mumpower, 1980; Howell & Burnett, 1978; Kubovy & Healy, 1980; Jungerman, Note 1). One of these led directly to the present experimentation: Howell & Burnett (1978) suggested that the source of the uncertainty and the specific nature of the required response constitute major task distinctions. More specifically, we contended that people form an impression of an event's likelihood based in part on what they believe is causing it and what, ultimately, they will have to do with it. Events over which they exercise some control, for example, tend to be perceived as more certain than the "objective" evidence justifies (Howell, 1971; Slovic, Fischhoff, & Lichtenstein, 1977); strongly held beliefs in some particular generative system tend to inhibit the processing of contra-

dictory evidence (Marques & Howell, Note 2); and anticipation of a data-based response such as frequency estimation could well produce better-calibrated impressions of uncertainty than do probability, prediction, or choice responses (Estes, 1976).

While the source-of-uncertainty hypothesis has received some empirical support as cited above, the response requirement predictions have yet to be adequately tested. The present research was designed to provide evidence on both distinctions within the context of the same realistic task setting. In the most general terms, it sought to determine whether people do indeed form substantially different impressions of uncertainty as a function of response requirement (frequency estimation, probability estimation, and predictive choice), and source of uncertainty (external vs. partially internal control of events). Since it is of course impossible to gain direct access to cognitions, the study took the form of a comparison of the "output" generated by subjects in response to comparable "objective inputs" under the various task conditions.

Before proceeding further with a description of the experiment, it would be well to consider some of the implications of the hypothesized task effects. If, as suggested above, task parameters help to shape the processes through which people construct intuitive uncertainty out of available evidence, then it is obviously important to know which are the influential task characteristics. Only in this way can we improve our ability to predict how people are likely to handle particular kinds of uncertain situations. However, a second, more pragmatic, implication is purely methodological. If one surveys the current research literature in almost any field of psychology with an



eye toward uncertainty concepts (expectation, confidence, subjective probability, perceived likelihood, etc.), he or she will find far greater apparent conceptual than operational agreement. The range of measures used to index this construct ranges from straight numerical expression to N-point rating scales to categorical judgments to various choice paradigms.<sup>1</sup> If, as we contend, different response requirements implicate different cognitive strategies -- at the encoding or retrieval stages of information processing, or both -- then there may be far less commonality in what is being measured than is typically assumed. It would thus seem useful to determine the extent of response-induced discrepancies within an otherwise common task framework.

Naturally, to undertake any evaluation of task variables requires the fixing of other task characteristics (to provide the "common task framework" referred to above). This, in turn, limits the generalizability of whatever results are obtained. In planning the present research, therefore, it seemed desirable to design a task scenario that would be representative of at least one important class of real-world situations for which people are required to form and use impressions of uncertainty. The defining features chosen were (a) frequentistic event base -- items to be processed occur repetitively, (b) complexity and processing load -- to-be-processed events are sufficient in number, variety, and presentation rate to obviate simple record-keeping (e.g., counting), (c) uncertainty task incidental -- ongoing task requirements permit, but do not emphasize, the formation of uncertainty impressions, and (d) plausibility or face validity -- subjects have reason to see the task as meaningful and worthy of

their continuing effort. The basic task developed to fulfill these requirements was that of a simulated resource-allocation position, dispatcher of coordinated emergency services, as described in the next section. The present research, therefore, was designed to determine whether people construct different impressions of the uncertainties inherent in realistic, complex, repetitive task settings as a function of event sources and response requirements.

#### METHOD

Task scenario. Subjects served individually in the role of dispatcher of emergency services for a hypothetical city. The area served was divided into 16 zones; the type of emergencies processed were of three types (police, fire, and ambulance calls). Thus any externally generated event (emergency) could be classified into one of 48 distinct categories. In addition, for reasons to be described shortly, any call could represent an actual emergency or a false alarm. Calls were programmed to occur at a rate dictated by the subject's response (but averaging about 2.5/min.) in daily "shifts" of about an hour's duration. The distribution of calls over the 96 event categories (48 x 2 levels of veracity) was as shown in Table 1 for each shift; the order of occurrence of the 150 calls presented

-----  
Table 1 about here  
-----

in a shift was randomized. Since the stochastic properties of the "generator" were stationary, subjects could acquire experience with the pattern of event uncertainties over shifts.

The subject's primary task was one of responding appropriately to each incoming call. Only two options were available: dispatching the required service immediately, or verifying the emergency (to rule out a false alarm). A cost/payoff scheme was devised in which verification became more desirable relative to immediate dispatching as the false alarm rate increased. Also, since the subject was allowed to distribute the limited emergency resources at the start of each shift, and there were costs associated with the distance between an available resource and each emergency, strategy played an important part in overall dispatching performance. In fact, a "score" was computed to reflect the quality of each decision; it was accumulated over an entire shift and displayed continuously to provide the subject with an indication of his ongoing performance.

The purpose of the entire allocation-decision scenario, of course, was simply to generate a variety of different frequentistic events to which subjects would be forced to attend in a setting that appeared realistic and intrinsically motivating. The actual decision performance was only of peripheral interest: it served chiefly as a method check to verify that subjects were indeed maintaining attention, that performance had asymptoted at the point of uncertainty measurement and that groups were comparable. Since it provided subjects with feedback on the appropriateness of each response (i.e., correct dispatch, dispatch to false alarm, verification of true emergency, verification of false alarm), the decision task also served as a source of internally controlled event frequencies. That is, a subject could be queried on the frequency of correct or incorrect responses

of various kinds (over which he had some control) just as he could on the frequency of specific types of calls (which were totally beyond his control).

The critical part of the task, from an experimental standpoint, came at the end of the fourth shift and throughout the fifth shift where all uncertainty measures were obtained. At the conclusion of the fourth regular shift (i.e., after 600 calls were processed) each subject was required to answer a series of questions regarding either the observed frequency of specified kinds of events (FE group) or the estimated probability of these same events in the future (PE group). These questions, of course, elicited all the desired uncertainty estimates for both externally generated events (types of calls) and internally controlled events (types of responses). Predictive choice measures of uncertainty were obtained during shift 5, the specific decision requests appearing individually, interleaved with the dispatching calls. Here, subjects were presented with a series of 25 pairs of events for each of which they were to choose the one more likely to occur next. The pairs were chosen to cover a wide range of past frequency differences and were presented one-at-a-time, without feedback, on the CRT display. Both the FE and PE groups performed the same series of choices as did a third (control, C) group which made neither type of estimation after the fourth shift.

The design, therefore, involved three groups (FE, PE and C) each of which was composed of 10 randomly assigned subjects all of whom observed the same distribution of emergency calls. However, between-group comparisons were of two kinds: direct estimation

scores, which reflected session 4 performance for PE and FE groups only; and choice scores, which reflected session 5 performance for all three groups. Comparisons involving the choice task thus indicated how prior estimation of either the frequency or probability of observed events affects subsequent decisions among these events. The PE - FE comparison, on the other hand, indicated how closely the two indexes of perceived uncertainty agree with each other and with "objective" values.

Subjects and Procedure. Subjects were recruited from several undergraduate psychology courses in which credit was awarded for experimental participation. In addition, bonuses were paid for performance on the choice task, the amounts averaging \$3.00 per subject. Assignment of subjects to the three groups was strictly random.

The experiment was carried out over five daily sessions (four dispatching shifts and one mixed dispatching-choice shift). Instructions were given during the initial session prior to the first shift and included an explanation of the purpose of the research ("to gain a better understanding of how people make resource-allocation decisions"), procedural instructions, and familiarization trials. Specific instructions for the estimation and choice tasks were given at the end of the fourth shift and the beginning of the fifth respectively.

The subject was seated in a small experimental booth before a TRS-80 micro-computer on the screen of which was displayed: (a) a map of the city zones, (b) a cumulative resource-allocation score for the shift, (c) an indication of available resources, (d) each emergency call as it appeared, (e) the response to each call as it

occurred, and (f) immediate feedback on the outcome of each response. The distribution and sequencing of the 150 events comprising each shift was programmed on tape. Subjects entered their responses to the calls using the computer keyboard, and the display of that call remained on the screen until the response was properly entered. Thus, the input sequence was largely self-paced (limited only by machine speed).

At the end of the fourth shift, the PE and FE questionnaires were administered to their respective groups. Items probed the uncertainty associated with specific event types, locations, location-type combinations, responses and correct responses. Examples of these items are illustrated in Table 2. The fifth session, as noted earlier,

-----  
Table 2 about here  
-----

consisted of both dispatching and predictive-choice trials for all three groups. The choice pairs were selected at random from all combinations of events having the same general character (i.e., event type, location, location-type combination) with the restriction that they represent the full available range of objective proportion differences (i.e., 1/1 through 6/1). It should be recognized that the limited sample of choice pairs, while fairly representative, could not possibly include representation from all combinations of ratios and event categories. Since our aim was simply to obtain an index of overall choice proficiency for use in comparing groups, not a precise psychometric function, this sample of pairs was deemed adequate.

At the end of the fifth session, bonuses were computed and paid, subjects were questioned informally about their strategies and general reaction to the task, a full explanation was given concerning the purposes of the research, and cooperation was solicited in maintaining the confidentiality of the true emphasis of the task. The reaction was uniformly favorable: subjects generally reported that the task was challenging and interesting, that they remained attentive throughout, and that their performance on the incidental uncertainty tasks far exceeded their expectations. Moreover, they contributed a number of useful suggestions for future improvement of the task. It would appear, therefore, that our primary task objectives were realized and that whatever criterion differences emerged could not reasonably be attributed to deficient or differential attention during the dispatching shifts.

#### RESULTS AND DISCUSSION

Method checks. In view of the complexity of the task scenario, it was important to establish that subjects were able to acquire proficiency on it, that their acquisition functions had stabilized by the time uncertainty measurement was undertaken, and that the three groups were comparable in both respects. A convenient index of this function was afforded by the allocation measure, a correlation of each subject's allocation of available resources with the optimum model for each shift. As shown in Figure 1, both acquisition

-----  
Figure 1 about here  
-----

and stabilization were verified. An analysis of variance indicated that the shifts effect was significant,  $F(4,27) = 70.87$  ( $p < .001$ ), but that neither group differences,  $F(2,27) = 1.95$  ( $p > .10$ ), nor the group  $\times$  shift interaction,  $F(8,108) < 1.0$ , approached significance. The fact that all groups achieved considerable proficiency also supports the contention that motivation was sustained over the five sessions.

Estimation. There are several ways to evaluate the quality of estimation performance, each of which has a somewhat different meaning. Since we are dealing exclusively with repetitive events generated according to a stable, stochastic process, there is in all cases a basis for an "objective" definition of frequency or probability. Thus, one can express performance in terms of unsigned deviations from this objective referent (calibration) or correlations with it (discrimination coefficient). The former, of course, combines "variable" with "constant" error, while the latter reflects the similarity of the profiles irrespective of any "constant" discrepancy. We analyzed individual performance using both measures, and the results showed large and consistent differences favoring the FE group regardless of the index used: the error scores averaged over subjects and items were 5.45 (FE) and 12.27 (PE); the correlations, .75 (FE) and .54 (PE). Analyses of variance supported this conclusion in both cases.

Considering first the unsigned deviation scores, both the group effect,  $F(1,18) = 5.69$ ,  $p < .03$ , and the event type effect,  $F(7,126) = 2.99$ ,  $p < .01$ , were statistically significant; the interaction, while just failing to achieve significance,  $F(7,126) = 1.99$ ,  $p = .06$ ,



cannot be summarily dismissed either. In this regard, it should be recognized that the questionnaire used to obtain estimates consisted of eight types of items, four (#1 - 4) probing externally generated events, and the other four (#5 - 8) events that were partially under the subject's control (response and response accuracy). Each of these categories was further subdivided into questions having a spatial component (#2, 4, 7, all of which probed 20 specific items), and non-spatial questions (#1, 3, 5, 6, 8, all of which necessarily probed only three or four items). To understand the meaning of the event type effect and the near-significant interaction, one must consider the pattern of unsigned error scores over these eight different kinds of questions. Presented in Table 3, these data show, first, that the

-----  
Table 3 about here  
-----

FE superiority was common to all types of items, and second, that it did not vary consistently with either the internal-external or the spatial-nonspatial categorizations. The main effect of event type is attributable, in large part, to the very good performance of both groups on questions #2 and #4 (the external spatial items), and the very poor performance on #5 (an internal non-spatial item). While internal items produced generally poorer estimates than did external items (average unsigned deviations of 11.12 vs. 6.60 respectively), our interpretation must be tempered by the fact that it was not possible to equate strictly the objective levels of these two classes of events. Still, this difference is consistent with earlier predictions (see Howell & Burnett, 1978) and with previous research

showing the existence of an "overconfidence" bias (Howell, 1971; Slovic, Fischhoff, & Lichtenstein, 1977).

The most noteworthy contrast in Table 2 involves the three 20-event spatial items (#2, 4, 7). Whereas the two external items (#2, 4) were handled very well by both groups, the internal item (#7) produced near-optimal estimates in the FE group and absolutely the worst estimates in the PE group. The nonsignificant interaction and similarly large FE - PE discrepancy for item #3 (estimated frequency or probability of false alarms by types) limits the interpretation that can safely be accorded this difference. However, the intriguing possibility it raises is that people form very accurate records of their own performance under some circumstances but rely upon entirely different information in projecting their future performance.

Turning to the correlation measure, it was possible to compute this index meaningfully only in the case of the three 20-event spatial items (#2, 4, 7). However, as shown in Table 4, the pattern

-----  
Table 4 about here  
-----

was much the same as for the error measure on these items. Overall, there was a distinct superiority of FE over PE performance; it was smallest for item #2 and largest for item #7. The group effect was highly significant,  $F(1,18) = 10.73$ ,  $p < .01$ , and the event-type effect -- here limited to two external and one internal category -- was only slightly less so,  $F(2,36) = 4.06$ ,  $p = .03$ . In this case, however, the interaction did not even approach significance, so not much should be made of the larger FE superiority on item #7.

The event-type effect on the correlation measure was not entirely consistent with that for the error index, largely because of the relative standing of items #4 and #7 on the two measures. Whereas #4 produced the lowest unsigned deviation scores (i.e., the highest accuracy) in both groups, it ranked lowest for both groups on the correlation index. One possible explanation for this anomaly is that subjects produced fewer extreme estimates on item #4, thereby reducing unsigned error but also restricting the range and depressing the correlation score. Another possibility is that the slightly higher error scores for item #2 and the markedly higher ones for #7 included a constant error component that was not present in item #4 performance; hence correlations were not as adversely affected as unsigned error scores. The latter explanation seems particularly plausible in the case of item #7 for the PE group, the only instance in which really poor error scores contrasted with rather moderate correlation scores. Since #7 was an internal item, a large constant error would be expected if people are, in fact, overconfident when their ability to control outcomes is involved; and the effect should be greatest when -- as in the PE group -- the estimation is directed toward anticipated future rather than demonstrated past performance (FE group). In any case, the absolute magnitude of the correlation coefficients for the FE group, even for as seemingly difficult a task as forming intuitive records of one's own performance on specific allocation problems (e.g., correctly dispatching ambulances to Zone 12), argues strongly for the viability of the experimental vehicle. Even the poorer estimates produced by the PE group were clearly influenced by the observed event frequencies; all  $r$ 's averaged well above zero, the differences

being highly significant:  $t(27) = 12.03, 9.37, \text{ and } 11.57$  for items #2, 4, and 7 respectively, all  $p < .001$ .

On the basis of the estimation performance, then, it seems safe to conclude that (a) although complex, the dispatching scenario affords subjects a reasonable opportunity to form data-based impressions of a variety of uncertainties, (b) probing stored frequency information yields consistently more vindical estimates than eliciting future-oriented probability judgments, (c) it is not clear that people form more accurate impressions of externally generated than internally generated repetitions, although there is a strong tendency for the probability estimation requirement to produce serious distortions of self-generated uncertainties -- possibly reflecting greater confidence than one's own stored record of past events should dictate, and (d) people seem particularly adept at organizing their intuitive records of event repetitions in spatial terms -- they have greater difficulty producing estimates by event types collapsed over locations.

Predictive-choice. It will be recalled that the comparison here was among three groups: those who had previously judged frequency (FE), probability (PE), and neither (C). The basis for comparison was performance on an identical set of 25 choice pairs administered during the fifth shift. Accuracy on these choice trials was defined in terms of probability differences. If prior experience in estimation alters the subject's mode of attacking decision problems, either by virtue of having formed a summary impression of various specific uncertainties or having formed the basis for a strategy to deal with uncertainties, it should be reflected in superior performance by FE

and PE over C. If, further, the greater accuracy produced by the FE orientation is incorporated into the summary or strategy used, that should be reflected in superior performance of FE over PE on the choice task.

As shown in Table 5, both of the above predictions received some support. Without the benefit of whatever cueing the estimation

-----  
Table 5 about here  
-----

task provided, accuracy for the C group was only slightly above chance (59%). Both estimation groups did substantially better than that, and the more accurately self-cued (FE) group was slightly superior (6%) to the PE group in average choice proficiency. A simple between-groups analysis of variance demonstrated the significance of the overall effect,  $F(2,27) = 3.60$ ,  $p < .05$ , although post-hoc analysis confirmed only the difference between C and the two estimation groups. In view of the limited criterion reliability afforded by a relatively small sample of choice pairs (25) plus the limited power afforded by only 10 subjects per group, one must consider these tests to be extremely conservative. Thus we regard the significant main effect as worthy of considerable attention, and the nonsignificant FE-PE difference of at least closer inspection in future research. In fact, the whole process by which estimation can influence subsequent choice, if indeed it does, would seem to justify more attention than it has so far received. From a theoretical standpoint, such information might help clarify discrepancies between "objective reality", judgment, and action in decision-making. From a practical standpoint,

it might suggest ways to promote better decision making. From a methodological standpoint, it might make results obtained using different uncertainty measures a bit easier to reconcile.

#### CONCLUSIONS

Viewed broadly, the present findings lend considerable support to the position that task characteristics, notably response requirements and types of events, have an important influence on uncertainty measures. Frequency estimation produced a considerably more accurate picture of repetitive events in a realistically complex setting than did probability estimation. Since both groups dealt with exactly the same "input" during the conditioning shifts, and indeed their allocation performance was strictly comparable, we must assume that the impressions formed were quite similar. The difference in obtained estimates must therefore be attributed to the incorporation of additional information in the "output" produced by the PE group. In other words, when asked to estimate the probability of specific events ("... between 0 - 100%, where 0% means you think the event has no chance of happening and 100% means you think it is certain . . . as in weather reports"), subjects apparently considered more than just their stored impression of past occurrences. What, precisely, the additional information was cannot be determined here, although there is a rather strong indication that an overconfidence bias may have been operating where the events were partially under the subject's control.

It comes as no great surprise that some kinds of event repetitions were stored better than others in this complex, dynamic

problem. Poor performance on items concerning the aggregate veracity of calls (false-alarm vs. truthful events) and reactions to them suggests that information is not normally encoded in this form. Likewise, generally good performance on both specific and general event patterns in the specific locations suggests that the spatial domain offers a rather natural framework for organizing repetitive occurrences. Since the present study was not designed to explore event coding dimensions in depth, these tentative conclusions deserve further investigation.

Estimating either the probability or frequency of repetitive events tends to improve the quality of subsequent choices involving those events. Although the evidence is inconclusive on this point, the amount of improvement seems to be related either to the estimation task or to the accuracy of estimates produced. Future research should address this phenomenon directly in order to discover how estimation influences decision performance. One possibility is that the act of verbalizing an estimate provides a handy memorial summary that is incorporated into the decision task. Another is that the estimation requirement merely cues information already in storage that serves to enhance decisions. Which process accounts for the effect is an empirical question that has, as noted earlier, a number of important practical and theoretical implications.

REFERENCE NOTES

- 1 Jungerman, H. Structural modeling of decision problems. Paper presented at American Psychological Association Meetings, Montreal Canada, Spt. 1-5, 1980.
2. Marques, T., & Howell, W.C. Intuitive frequency judgments as a function of prior expectations, observed evidence, and individual processing strategies. Technical Report No. 79-06, Department of Psychology Report Series, Houston: Rice University, 1979.



## REFERENCES

- Cohen, L.J. On the psychology of prediction: Whose is the fallacy? Cognition, 1979, 7, 385-407.
- Dawes, R.M., & Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.
- Einhorn, H.J., & Hogarth, R.M. Behavioral decision theory: processes of judgment and choice. Annual Review of Psychology, 1981, 32, 53-88.
- Estes, W.K. The cognitive side of probability learning. Psychological Review, 1976, 83, 37-64.
- Hammond, K.R., McClelland, G.H., & Mumpower, J. Human judgment and decision making: theories, methods, and procedures. New York: Praeger, 1980.
- Howell, W.C. Uncertainty from internal and external sources: a clear case of overconfidence. Journal of Experimental Psychology, 1971, 89, 240-243.
- Howell, W.C., & Burnett, S.A. Uncertainty measurement: A cognitive taxonomy. Organizational Behavior and Human Performance, 1978, 22, 45-68.
- Kahneman, D., & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 251-273.
- Kahneman, D., & Tversky, A. Prospect theory: an analysis of decision under risk. Econometrica, 1979, 47, 263-291.
- Kubovy, M., & Healy, A.F. Process models of probabilistic categorization. In T.S. Wallsten (Ed.). Cognitive processes in choice and decision behavior. Hillsdale, New Jersey: Erlbaum, 1980.

- Lichtenstein, S., Fischhoff, B., & Phillips, L.D. Calibration of probabilities: the state of the art. In H. Jungermann & G. de Zeeuw (Eds.), Proceedings of the fifth research conference on subjective probability, utility, and decision making. Dordrecht, the Netherlands: D. Reidel, 1977.
- Naylor, J.C., and Domine, R.K. Inferences based on uncertain data: some experiments on the role of slope magnitude, instructions, and stimulus distribution shape on the learning of contingency relationships. Organizational Behavior and Human Performance, 1981, 27, 1-13.
- Slovic, P., Fischhoff, B. & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.
- Slovic, S. & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.
- Tversky, A., & Kahneman, D. Judgment under uncertainty: heuristics and biases. Science, 1974, 185, 1124-1131.
- Tversky, A., & Kahneman, D. The framing of decisions and the psychology of choice. Science, 1981, 211, 453-458.
- Wallsten, T.S. (Ed.) Cognitive processes in choice and decision behavior. Hillsdale, New Jersey: Erlbaum, 1980.

## FOOTNOTES

<sup>1</sup>We did, in fact, conduct such a survey on a very informal basis as reported by the first author at the 1980 American Psychological Association Meetings in Montreal Sept. 1-5, 1980. The point is sufficiently obvious not to justify specific citations here.

TABLE 1

Distribution of calls over the 96 event categories  
classified by location, type of emergency and level  
of veracity

Location	Type of Emergency					
	<u>Police</u>		<u>Fire</u>		<u>Ambulance</u>	
	AE	FA	AE	FA	AE	FA
1	2	1	-	-	2	-
2	1	-	4	1	-	-
3	1	-	-	-	-	-
4	-	-	-	-	-	-
5	10	3	-	-	4	1
6	8	2	-	-	9	2
7	7	2	-	-	4	-
8	10	5	3	-	6	-
9	2	2	2	-	7	3
10	-	-	7	1	10	1
11	10	2	-	-	1	-
12	3	4	1	-	-	-
13	-	-	-	-	-	-
14	-	-	1	1	-	-
15	-	-	-	-	-	-
16	-	-	4	-	-	-

Note: AE = actual emergency; FA = false alarm

TABLE 2

The Eight Categories of Information Probed with Illustrations of Comparable Items Administered to the Two groups (FE and PE).

<u>Probe Category</u>	<u>FE Group</u>	<u>Example*</u>	<u>PE Group</u>
1. Type of event (police, fire, amb.)	How many total <u>police</u> calls did you receive?		If a call comes in, what are the chances (0-100%) that it will be a <u>police</u> call?
2. Type by location	(Map presented with instructions to estimate the totals for events indicated), e.g., <u>police calls, zone 8.</u>		(Map presented with instructions to estimate the chances, 0-100%, for events indicated), e.g., <u>police call, zone 8.</u>
3. Type by veracity	How many false alarms were ( <u>police calls</u> )?		Suppose a call was a false alarm. What are the chances of its being a ( <u>police</u> ) call?
4. Type by location by veracity	(Map presented.) Please fill in the totals for false alarms only.		(Map presented.) Again, suppose a call was a false alarm. What are the chances it would be of the type and location indicated?
5. Response to veracity of calls	On how many occasions did you <u>verify a false alarm</u> ?		For any given call, what are the chances that you would <u>verify a false alarm</u> ?
6. Response to event type	How often did you correctly or incorrectly dispatch a ( <u>police</u> ) unit?		If you dispatched a unit, correctly or incorrectly, what are the chances that it was a ( <u>police</u> ) unit?
7. Correct response to event type by location	(Map presented with instructions to estimate the number of units dispatched correctly for indicated type/location.)		(Map presented with instructions to estimate the chances that a unit correctly dispatched would be of indicated type/location.)
8. Correct response to event type	How often did you correctly dispatch a ( <u>police</u> ) unit?		If you correctly dispatched a unit, what are the chances that it was a ( <u>police</u> ) unit?

\*These examples are designed to give the reader a general idea of the probes used, not the exact format or total context of the questionnaire. For example, considerable explanation was given, replete with illustrations, to insure the subject's understanding of the PE and FE response concepts.

TABLE 3  
Mean Unsigned Error Scores for the Two  
Estimation Groups over the Eight  
Probe Categories

<u>Probe Category</u>	<u>FE</u>	<u>Group</u> <u>FE</u>
1. type of event	6.03	7.60
2. type by location	2.93	7.16
3. type by veracity	4.20	16.71
4. type by location by veracity	.98	7.17
5. response to veracity of calls	11.66	16.02
6. response to event type	9.93	12.61
7. correct response to event type by location	1.45	18.83
8. correct response to event type	6.38	12.04

TABLE 4

Mean Correlations\* for the Frequency (FE) and  
Probability (PE) Estimation Groups over the  
Three Spatial Probes (#2, 4, 7)

Group	<u>Probe Category**</u>		
	2. Type by location	4. Type by Location by veracity	7. Correct res- ponse to event type by location
FE	.78	.66	.82
PE	.64	.44	.54

\*n = 20

\*\* Note: Items #2 & 4 probe externally generated events, and #7  
probes events partially under the subjects' control.

TABLE 5

Predictive Choice Accuracy and Overall  
Estimation Performance for the Three  
Groups

<u>Group</u>	<u>Percentage Correct Choices</u>	<u>Average Estimates</u>	
		<u>Correlation</u>	<u>Unsigned Error</u>
FE	77.6	.75	5.45
PE	71.6	.54	12.27
Control	59.2	--	---



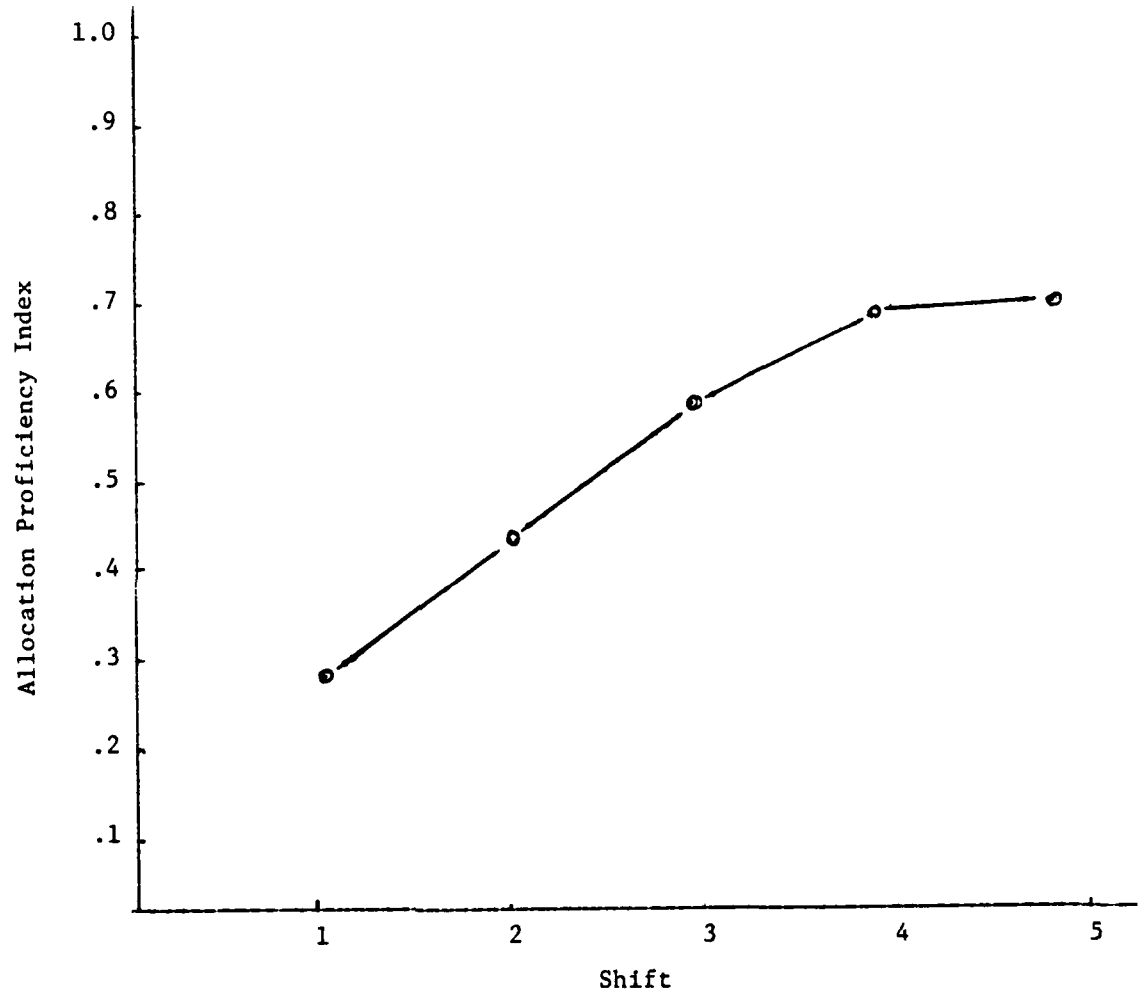


Figure 1. Acquisition Function for Allocation Performance Averaged Over Groups.

N = 30

OFFICE OF NAVAL RESEARCH

Code 455

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CDR Paul R. Chatelier  
Office of the Deputy Under  
Secretary of Defense  
OUSDRE (E&LS)  
Pentagon, Room 3D129  
Washington, DC 20301

Department of the Navy

Director  
Engineering Psychology Programs  
Code 455  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217 (5 cys)

Director  
Operations Research Programs  
Code 434  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217

Director  
Statistics and Probability Program  
Code 436  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217

Director  
Information Systems Program  
Code 437  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217

Code 430B  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217

Department of the Navy

Special Assistant for Marine Corps  
Matters  
Code 100M  
Office of Naval Research  
800 North Quincy Street  
Arlington, VA 22217

Commanding Officer  
ONR Eastern/Central Regional Office  
ATTN: Dr. J. Lester  
Building 114, Section D  
666 Summer Street  
Boston, MA 02210

Commanding Officer  
ONR Branch Office  
ATTN: Dr. C. Davis  
536 South Clark Street  
Chicago, IL 60605

Commanding Officer  
ONR Western Regional Office  
ATTN: Mr. R. Lawson  
1030 East Green Street  
Pasadena, CA 91106

Commanding Officer  
ONR Western Regional Office  
ATTN: Dr. E. Gloye  
1030 East Green Street  
Pasadena, CA 91106

Office of Naval Research  
Scientific Liaison Group  
American Embassy, Room A-407  
APO San Francisco, CA 6503

Director  
Naval Research Laboratory  
Technical Information Division  
Code 2627  
Washington, D.C. 20375 (6 cys)

Department of the Navy

Dr. Bruce Wald  
Communications Sciences Division  
Code 7500  
Naval Research Laboratory  
Washington, D.C. 20375

Dr. Robert G. Smith  
Office of the Chief of Naval  
Operations, OP987H  
Personnel Logistics Plans  
Washington, D.C. 20350

Naval Training Equipment Center  
ATTN: Technical Library  
Orlando, FL 32813

Human Factors Department  
Code N215  
Naval Training Equipment Center  
Orlando, FL 32813

Dr. Alfred F. Smode  
Training Analysis and Evaluation  
Group  
Naval Training Equipment Center  
Code N-00T  
Orlando, FL 32813

CDR Thomas Berhage  
Naval Health Research Center  
San Diego, CA 92152

Dr. George Moeller  
Human Factors Engineering Branch  
Submarine Medical Research Lab  
Naval Submarine Base  
Groton, CT 06340

Dr. James McGrath, Code 302  
Navy Personnel Research and  
Development Center  
San Diego, CA 92152

Navy Personnel Research and  
Development Center  
Planning & Appraisal  
Code 04  
San Diego, CA 92152

Department of the Navy

Navy Personnel Research and Develop-  
ment Center  
Management Systems, Code 303  
San Diego, CA 92152

Navy Personnel Research and Develop-  
ment Center  
Performance Measurement & Enhancement  
Code 309  
San Diego, CA 92152

CDR P. M. Curran  
Code 604  
Human Factors Engineering Division  
Naval Air Development Center  
Warminster, PA 18974

Mr. Ronald A. Erickson  
Human Factors Branch  
Code 3194  
Naval Weapons Center  
China Lake, CA 93555

Dean of the Academic Departments  
U.S. Naval Academy  
Annapolis, MD 21402

Dr. Gary Poock  
Operations Research Department  
Naval Postgraduate School  
Monterey, CA 93940

Dean of Research Administration  
Naval Postgraduate School  
Monterey, CA 93940

Mr. Warren Lewis  
Human Engineering Branch  
Code 8231  
Naval Ocean Systems Center  
San Diego, CA 92152

Dr. A. L. Slafkosky  
Scientific Advisor  
Commandant of the Marine Corps  
Code RD-1  
Washington, D.C. 20380

Department of the Navy

Mr. Arnold Rubinstein  
Naval Material Command  
NAVMAT 0722 - Rm. 508  
800 North Quincy Street  
Arlington, VA 22217

Mr. Phillip Andrews  
Naval Sea Systems Command  
NAVSEA 0341  
Washington, D.C. 20362

Commander  
Naval Electronics Systems Command  
Human Factors Engineering Branch  
Code 4701  
Washington, D.C. 20360

LCDR W. Moroney  
Code 55MP  
Naval Postgraduate School  
Monterey, CA 93940

Mr. Merlin Malehorn  
Office of the Chief of Naval  
Operations (OP-115)  
Washington, D.C. 20350

Department of the Army

Mr. J. Barber  
HQS, Department of the Army  
DAPE-MBR  
Washington, D.C. 20310

Dr. Joseph Zeidner  
Technical Director  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Director, Organizations and  
Systems Research Laboratory  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Technical Director  
U.S. Army Human Engineering Labs  
Aberdeen Proving Ground, MD 21005

U.S. Army Medical R&D Command  
ATTN: DPT Gerald P. Krueger  
Ft. Detrick, MD 21701

Department of the Army

ARI Field Unit - USAREUR  
ATTN: Library  
C/O ODCSPER  
HQ USAREUR & 7th Army  
APO New York 09403

Department of the Air Force

U.S. Air Force Office of Scientific  
Research  
Life Sciences Directorate, NL  
Bolling Air Force Base  
Washington, D.C. 20332

Dr. Donald A. Topmiller  
Chief, Systems Engineering Branch  
Human Engineering Division  
USAF AMRL/HES  
Wright-Patterson AFB, OH 45433

Air University Library  
Maxwell Air Force Base, AL 36112

Dr. Earl Alluisi  
Chief Scientist  
AFHRL/CCN  
Brooks AFB, TX 78235

Foreign Addresses

North East London Polytechnic  
The Charles Myers Library  
Livingstone Road  
Stratford  
London E15 2LJ  
ENGLAND

Professor Dr. Carl Graf Hoyos  
Institute for Psychology  
Technical University  
8000 Munich  
Arcisstr 21  
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner  
Applied Psychology Unit  
Admiralty Marine Technology Establishment  
Teddington, Middlesex TW11 0LN  
ENGLAND

Foreign Addresses

Director, Human Factors Wing  
Defense & Civil Institute of  
Environmental Medicine  
Post Office Box 2000  
Downsview, Ontario M3M 3B9  
CANADA

Dr. A.D. Baddeley  
Director, Applied Psychology Unit  
Medical Research Council  
15 Chaucer Road  
Cambridge, CB2 2EF  
ENGLAND

Other Government Agencies

Defense Technical Information  
Center  
Cameron Station, Bldg. 5  
Alexandria, VA 22314 (12 cys)

Dr. Craig Fields  
Director, Cybernetics Technology  
Office  
Defense Advanced Research Projects  
Agency  
1400 Wilson Blvd.  
Arlington, VA 22209

Dr. Judith Daly  
Cybernetics Technology Office  
Defense Advanced Research Projects  
Agency  
1400 Wilson Blvd.  
Arlington, VA 22209

Professor Douglas E. Hunter  
Defense Intelligence School  
Washington, D.C. 20374

Other Organizations

Dr. Robert R. Mackie  
Human Factors Research, Inc.  
5775 Dawson Avenue  
Goleta, CA 93017

Dr. Gary McClelland  
Institute of Behavioral Sciences  
University of Colorado  
Boulder, CO 80309

Other Organizations

Dr. Miley Merkhofer  
Stanford Research Institute  
Decision Analysis Group  
Menlo Park, CA 94025

Dr. Jesse Orlansky  
Institute for Defense Analyses  
400 Army-Navy Drive  
Arlington, VA 22202

Professor Judea Pearl  
Engineering Systems Department  
University of California - Los Angeles  
405 Hilgard Avenue  
Los Angeles, CA 90024

Professor Howard Raiffa  
Graduate School of Business Admin-  
istration  
Harvard University  
Soldiers Field Road  
Boston, MA 02163

Dr. Arthur I. Siegel  
Applied Psychological Services, Inc.  
404 East Lancaster Street  
Wayne, PA 19087

Dr. Paul Slovic  
Decision Research  
1201 Oak Street  
Eugene, OR 97401

Dr. Amos Tversky  
Department of Psychology  
Stanford University  
Stanford, CA 94305

Dr. Robert T. Hennessy  
NAS - National Research Council  
JH #819  
2101 Constitution Ave., N.W.  
Washington, DC 20418

Dr. M. G. Samet  
Perceptronic, Inc.  
6271 Variel Avenue  
Woodland Hills, CA 91364

Other Organizations

Dr. Robert Williges  
Human Factors Laboratory  
Virginia Polytechnical Institute  
and State University  
130 Whittemore Hall  
Blacksburg, VA 24061

Dr. Alphonse Chapanis  
Department of Psychology  
The Johns Hopkins University  
Charles and 34th Streets  
Baltimore, MD 21218

Dr. Meredith P. Crawford  
American Psychological Association  
Office of Educational Affairs  
1200 17th Street, NW  
Washington, DC 20036

Dr. Ward Edwards  
Director, Social Science Research  
Institute  
University of Southern California  
Los Angeles, CA 90007

Dr. Charles Gettys  
Department of Psychology  
University of Oklahoma  
455 West Lindsey  
Norman, OK 73069

Dr. Kenneth Hammond  
Institute of Behavioral Science  
University of Colorado  
Room 201  
Boulder, CO 80309

Journal Supplement Abstract Service  
American Psychological Association  
1200 17th Street, NW  
Washington, DC 20036 (3 cys)

Dr. John Payne  
Duke University  
Graduate School of Business Admin-  
istration  
Durham, NC 27706

Dr. Baruch Fischhoff  
Decision Research  
1201 Oak Street  
Eugene, OR 97401

Other Organizations

Dr. Andrew P. Sage  
University of Virginia  
School of Engineering and Applied  
Science  
Charlottesville, VA 22901

Dr. Leonard Adelman  
Decisions and Designs, Inc.  
8400 Westpark Drive, Suite 600  
P.O. Box 907  
McLean, VA 22101

Dr. Lola Lopes  
Department of Psychology  
University of Wisconsin  
Madison, WI 53706

DATE  
FILMED  
-8